DOCUMENT RESUME

ED 262 476 EA 018 028

AUTHOR Greene, Jennifer C.
TITLE Qualitative Data Analysis in Program Evaluation Contexts.
PUB DATE 85
NOTE 39p.; Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, IL, March 31-April 4, 1985).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Case Studies; *Data Analysis; Guidelines; *Program Evaluation; Research Design; *Research Methodology
IDENTIFIERS *Qualitative Analysis; *Qualitative Data

ABSTRACT

The naturalistic research perspective assumes that reality is multiplistic, phenomenological, and context-dependent. This perspective legitimizes the subjective insights of the investigator by acknowledging the interdependence of facts and values as well as of the investigator and the object of investigation. Although discrepancies between investigators' interpretations of given phenomena are not problematical for naturalistic research in general, they are problematical when the naturalistic perspective is applied to program evaluation. This document presents a framework for developing qualitative data analysis procedures for program evaluation that will lead to results that can be broadly, convincingly, and credibly defended. Three naturalistic studies using such procedures are described to illustrate the application of the framework and the effectiveness of the procedural attitude recommended. A three-page list of references and supplemental tables complete the report. (PGD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Jennifer C. Greene

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED262476

Qualitative Data Analysis in Program Evaluation Contexts

Jennifer C. Greene

Cornell University

Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, 1985.

EA 018 028

Abstract

Toward the "demystification" of qualitative data analysis within the naturalistic perspective on inquiry, examples of and reflections on the analysis strategies used in three naturalistic studies are shared. These examples include strategies used during and after data collection, as well as in conjunction with initial results and interpretations. The focus throughout is on qualitative analyses in program evaluation contexts, with the aim of strengthening the acceptability of and thus opportunities for using naturalistic approaches to evaluation.

Qualitative Data Analysis in Program Evaluation Contexts

Data analysis is the process of reducing, sorting, and organizing a set of raw data into units, clusters, and/or patterns. The function of data analysis is to provide the basis for interpretation and inference or the process of attaching meaning, explanation, and significance to the units, clusters, and patterns identified. When the raw data are numbers, data analysis is usually a matter of selecting and correctly implementing the appropriate statistical procedures. When the raw data are words, however -- fieldnotes, interview transcripts, documents and records -- the process of data analysis is less easily codified. "Words are fatter than numbers, and usually have multiple meanings. This makes them harder to move around and work with" (Miles & Huberman, 1984, p. 54).

Yet, the generation of raw data as words is increasingly common in program evaluation studies, reflecting the recent, widespread acceptance of qualitative methods within the evaluation community. Along with this acceptance have come important reference tools to help the "would-be qualitative evaluator" (who, until recently, is typically well schooled only in quantitative methods) learn about the underpinnings and implementation of qualitative methodology (e.g., Bogdan & Biklen, 1982; Fetterman, 1984; Guba & Lincoln, 1981; Miles & Huberman, 1984; Patton, 1980). These references all address the problems and processes of qualitative data analysis, though with varying degrees of specificity. More to the point of the present paper,

these and other references discuss qualitative data analysis with varying degrees of adherence to the epistemological assumptions commonly linked with qualitative methodology.[1]

These epistemological assumptions are embodied in the naturalistic perspective on inquiry as presented by Guba and Lincoln (1981).[2] From this perspective, complete a priori specification of data collection and analysis procedures for a given study is neither possible nor desirable because the inherent nature of naturalistic inquiry is emergent, grounded, and context-bound. Naturalistic data analysis is intertwined with both data collection and interpretation throughout the process of inquiry. It also relies heavily on the inductive insights of the evaluator as s/he continuously generates, empirically checks out, and refines "working hypotheses" relevant to the inquiry. Thus, because procedural prescriptions for data analysis are incompatible with this perspective, the process of analysis is an ongoing challenge.

The Miles and Huberman (1984) "sourcebook" of qualitative data analysis methods represents an important response to this challenge, arising from the authors' own experience with the "mysterious, half-formulated ... arcane process of making valid sense of large amounts of qualitative information" (Miles, 1983, p. 122 and 125). Drawing on this experience, the authors present in this sourcebook a varied and creative array of systematic methods for analyzing, displaying, and interpreting qualitative data, including many inventive uses of matrices and other graphic

devices. However, their clearly stated stance in this book is one of "soft-nosed logical positivism," representing a "middle-of-the-road," eclectic blend of a valuing of formalized, verifiable methods as requisite for valid conclusions and a valuing of inductive phenomenology as useful for illuminating social phenomena. This stance is reflected in their repeated emphasis on variables, causal modelling, and replication. This stance is also the target of critiques that caution sourcebook readers about undermining the essential purposes and strengths of naturalistic inquiry and that chide Miles and Huberman for discounting as "mysterious and arcane" the existing canons of qualitative data analysis (Marshall, 1984).

The stance taken in this paper is "middle-of-the-road" between Miles and Huberman and their critics. I agree with Miles and Huberman that the process of qualitative data analysis needs "demystification," and I find their sourcebook to be a stimulating collection of analytic ideas. I also agree with some critics' laments that the positivist criteria underlying Miles and Huberman's concept of validity endanger the emic perception of meaning that is the heart of naturalistic inquiry. In short, I believe we need further specification of concrete, practical qualitative data analysis procedures for use <u>within</u> the naturalistic approach to inquiry.

One important vehicle for addressing this need is to encourage naturalistic evaluators and researchers to share the specific analytic methods they have used, along with their

reflections on the benefits and limitations of these methods (Alberty in Patton, 1980; Daillak & Alkin, 1981; Donmoyer, 1984; Greene et al., 1984; and Miles & Huberman, 1984). The present paper is offered in this growing tradition. More specifically, a review of general methodological guidance on qualitative data analysis will provide the framework for a discussion of the analytic strategies used in three different naturalistic studies. The emphasis throughout will be on qualitative data analysis within naturalistic evaluative inquiry. A brief overview of the latter will first provide the relevant context for this discussion.

Naturalistic Program Evaluation

All approaches to program evaluation seek to provide information about one or more aspects of a program that is useful for one or more designated audiences. Differences among evaluation approaches arise primarily from their differential orientations to the program information needs of varying audiences and, secondarily, from the inquiry perspective and methodology most compatible with a given orientation. For example, from House (1981), a "behavioral objectives" approach to evaluation focuses on program administrators' needs for program outcome and accountability information. In contrast, a "case study" approach emphasizes staff and clients' needs for information about the diversity of program experiences. And though these different orientations may suggest different methodologies (e.g., using a quasi-experimental vs. an

ethnographic perspective, respectively), this latter linkage is neither automatic nor inherent. [See, for example, Campbell's (1979) discussion of degrees of freedom within a single case study.]

So, a naturalistic program evaluation is not really distinguished by the program aspect studied (e.g., implementation vs. outcomes) nor by the audiences served. Rather, guided by underlying epistemological assumptions, its most important distinctiveness comes in the way "information" (or what is important to know) is conceptualized.

In brief, the naturalistic perspective assumes that reality is multiplistic, phenomenological, and context-dependent and thus that understanding human phenomena requires an understanding of the multiple ways individuals construct reality and find meaning for themselves within one or more identified contexts. The naturalistic perspective further acknowledges the interdependence of facts and values and of the investigator and the object of the investigation and, in doing so, legitimizes the individual, "subjective" insights of the investigator. Hence, a naturalistic program evaluator will seek to offer his/her own descriptive account and interpretation of the ways in which multiple, relevant program constituencies perceive, understand, and find meaning in their differing program experiences (with an appropriate focus on the specific evaluation questions of interest). Further, given the important legitimization of the evaluator's own insights, a different evaluator may very well

generate a different set of results and interpretations.

That is, differences or even discrepancies between investigators' interpretive understandings of a given phenomena are not particularly problematic for the naturalistic perspective itself. However, as will be argued more fully later in this paper, the possibility of such discrepancies is problematic for naturalistic program evaluation, or the use of the naturalistic perspective in program evaluation contexts. The essence of this argument is that the potentially impactive, political contexts of program evaluation render unacceptable evaluation results accompanied by "This is just my view." Required instead are evaluation results that can be more broadly, convincingly, and credibly defended.

A framework of qualitative data analysis procedures for attaining such results is presented in the next section, followed by analysis examples and reflections thereon.

General Methodological Guidance for Qualitative Data Analysis

Within the contemporary context of naturalistic program evaluation, the existing canons of qualitative data analysis can be roughly grouped into three clusters according to their prominence during different inquiry phases.[3] (a) Most salient during data collection are processes related to analytic induction. (b) Upon completion of data collection, the processes of inductive content analysis are commonly most relevant. (c) Then, along with a preliminary summary of findings come the processes of checking and auditing these findings for trustworthiness.

Brief descriptions of each cluster are presented below.

Data analysis during data collection: Processes related to analytic induction. The cyclical intertwining of naturalistic data collection, analysis, and interpretation is captured in the key naturalistic concept of analytic induction. Analytic induction refers to the ongoing, interrelated processes of sorting and organizing the data collected thus far, using this organization to generate working hypotheses or speculations about possible patterns of meaning, and determining which "next steps" in data collection will most effectively help to check out, test, or augment these preliminary hypotheses (Becker, 1970; Bogdan & Biklen, 1982; Denzin, 1978). These ongoing data collection decisions typically include theoretical or purposive sampling, negative case identification and analysis, and data and methodological triangulation (Denzin, 1978; Guba & Lincoln, 1981; Patton, 1980).

But, how is the analysis part of analytic induction conducted or how are existing fieldnotes, interview transcripts, and/or documents reduced, sorted, and organized into manageable chunks for preliminary interpretation? It is in response to the apparent "mystery" of this process that Miles and Huberman (1984) compiled their sourcebook. Yet, while their sourcebook offers numerous procedural aids for qualitative analysis during data collection, the essence of analytic induction lies within the individual analyst's processes of perception, insight, and reasoning. As such, it will always remain at least partially

hidden, or mysterious if you will, in contrast to the fully explicated, step-by-step procedures for conducting an ANOVA or a multiple regression. Moreover, within the naturalistic perspective, this "intuitive" data processing is of central value and importance, i.e., organizing data as a vehicle for understanding an emic perspective requires the insight and reasoning of another human being. Further, with this reliance on insight and reasoning comes the acknowledged importance of constant monitoring for evaluator biases or preconceptions (Guba & Lincoln, 1981), for evaluator effects (Guba & Lincoln, 1981; Patton, 1980), and for "information processing" errors (Sadler, 1981).[4]

This perspective on Miles and Huberman's presentation notwithstanding, their procedural aids for qualitative analysis during inquiry are appropriate and helpful. These aids can be grouped into several categories (see also Bogdan & Biklen, 1982; Gillespie, 1982; Glaser, 1978; Patton, 1980).

1. Include a set of analytic comments as part of each data collection log or record, e.g., observer comments, analytic questions, reflective or marginal remarks, working hypotheses.

2. Write analytic "memos," as periodic syntheses of the analytic comments in individual data logs and as documentation of sudden insights or developing conceptual understandings.

3. Consult periodically with project team members and other colleagues about emerging hypotheses, e.g., in a site analysis meeting or a peer debriefing.

4. Develop and apply codes for data as a means of data

reduction and as a vehicle for pattern identification. (See discussion below on inductive content analysis.)

5. Adopt an attitude of adventurous speculation, e.g., through the use of metaphors or analogies, remembering that while working hypotheses are not final conclusions, such preliminary, speculative understandings underlie subsequent conclusions that are defensible and credible. This last point underscores the fundamental importance of analytic induction as a cornerstone of qualitative inquiry.

Data analysis after data collection: The process of inductive content analysis. As an inquiry technique, content analysis has centuries-old history in the field of communications and has been most widely applied to existing documents, records, and such other existing qualitative materials as newspaper articles and advertisements (Berelson, 1952; Holsti; 1969; Krippendorf, 1980). As a data analysis strategy within the naturalistic approach to inquiry, inductive content analysis shares some important characteristics with its origins, but also diverges in several other key aspects, as shown in Table 1. The difference in terminology is thus a deliberate attempt to stem continuing confusion regarding the meaning, intent, and implementation of inductive content analysis within the naturalistic paradigm.

Insert Table 1 about here

Inductive content analysis then refers to the interrelated and fully iterative processes of (a) sorting qualitative data into categories; (b) organizing, labelling, and revising these categories; (c) re-sorting by applying the categories back to the data (i.e., coding); (d) using this re-sort to reorganize, relabel, and revise again the categories; and so forth. The goal of this analysis strategy is a set of categories that provides a comprehensive and meaningful framework for summarizing the raw data. This strategy requires prior decisions about units of analysis, as well as a set of final categories that meets the specifications listed in Table 1. Thus, for a complex, multidimensional study, several inductive content analyses may be needed, each focused on a different conceptual unit of analysis and/or each based on a different classification principle.

Procedural suggestions for sorting and classifying data emphasize searching for natural variations or "recurring regularities," e.g., by marking or writing these down while reading through the data several times or by physically cutting up a copy of the data into analysis units and then physically sorting these units into similar piles (Bogdan & Biklen, 1982; Patton, 1980). While the conceptual nature or type of category system(s) to be developed should reflect the purpose of inquiry, many authors have offered suggestions and examples. Lofland (1971), for example, suggests that potential phenomena for categorization and coding range from brief acts and ongoing activities to multiple relationships among people and entire

settings. (See Bogdan & Biklen, 1982; Guba & Lincoln, 1981; Miles & Huberman, 1984; and Patton, 1980 for additional suggestions and examples.)

Thus, within naturalistic inquiry, inductive content analysis serves primarily as a data reduction and organization or classification technique that provides the basis for interpretation and inference. Clearly, some interpretation will inevitably accompany this analytic endeavor, just as continuing analysis (e.g., refinements to the category system) will inevitably accompany the summarization and interpretation of results. This simply reemphasizes the fully intertwined nature of naturalistic data collection, analysis, and interpretation. Additional analytic strategies accompanying results summaries and interpretations are discussed next.

Data analysis for assessing the trustworthiness of summary findings and interpretations. Naturalistic inquiry has its own analogs to covariate controls, reliability coefficients, and statistical significance in the form of procedures for assessing the confirmability, dependability, and credibility of results (Guba & Lincoln, 1981). Many of these procedures, some noted above, are integral parts of data collection and ongoing analysis (e.g., triangulation, purposive sampling, negative case analysis, reflexivity). Others are more appropriately implemented in conjunction with a "first draft" of summary findings and interpretations. They are still considered analysis strategies in that they relate to the quality of the data base underlying these

results. Key examples of these strategies are listed below.

1. Commission external "audits" of (a) the degree to which the summary results are supported by the available data base and (b) the logical consistency and soundness of the methodological decisions made throughout the inquiry process. In Guba and Lincoln's (1981) terminology, these are a confirmability and a dependability audit, respectively. Miles and Huberman (1984) provide useful suggestions for how to maintain the information needed for such audits (the "audit trail"), and Schwandt (1984) offers an instructive example of both.

2. Check out the credibility of descriptive summary results with members of the setting being studied. These member checks, which can be conducted with both participants and nonparticipants in the inquiry process, are a vital source of results confirmation. This form of member checks focuses on the credibility of the evaluator's descriptive summary of the data collected, not his/her interpretation thereof. Thus, confirmation (e.g., "Yes, that's what I told you" or "Yes, that's what the annual report says") can be expected.

3. Also check out rival interpretations of descriptive results via critical reflection, peer debriefings, member checks again, as well as discussion of such alternatives in a summary report. For member checks, this form focuses on the credibility of the evaluator's interpretation of descriptive findings. Thus, alternative interpretations rather than confirmation can be expected, because such multiplicity is integral to the fabric of

naturalistic inquiry. Moreover, as illustrated by Donmoyer (1983), these multiple interpretations can be woven into the tapestry of a summary report, within which the evaluator can offer his/her preferred interpretation along with its rationale. (For example, Donmoyer's (1984) suggested rationale for "choosing among plausible alternatives" is to choose the one that is most congruent with the purpose of the inquiry.)

While intuitively appealing to those of us concerned about methodological rigor, these analytic suggestions for assessing the trustworthiness of naturalistic results have recently been challenged by Smith (1984). Smith (1983a, 1983b; Smith & Heshusius, 1985) has been a persistent and eloquent spokesperson for the importance of maintaining the integrity of the naturalistic (interpretive) perspective. In brief, his recent argument is that the very existence of "foundational criteria" or standards for assessing trustworthiness is incompatible with the assumptions and purposes of naturalistic inquiry. The very notion of naturalistic analogs to reliability and validity is epistemologically inconsistent, a philosophical oxymoron, if you will.

> To accept that social reality is mind-constructed and that there are multiple realities is to deny that there are any "givens" upon which to found knowledge. If one accepts these assumptions, different claims about reality result not from incorrect procedures but may simply be a case of one investigator's interpretation

> of reality versus another's. In a world of multiply constructed, mind-dependent realities, there may be no "court of last resort" to appeal to to sort out trustworthy interpretations from untrustworthy ones. (1984, p. 383)

While not venturing to counter Smith's argument in the philosophical or epistemological arenas, I do wish to contend that trustworthiness analyses are important for the practical arena of program evaluation. I would like to suggest further that the use of such analyses need not necessarily violate the assumptions of naturalistic inquiry.

First, the practical importance of trustworthiness assessments for program evaluation is underscored by the "real world" setting of evaluative inquiry. Despite our continuing concerns about underutilization, as well as our more modest expectations about potential utility, evaluation results can and do make a difference in people's lives. This difference could be an improved program activity for participants, greater understanding of audience needs or program delivery dynamics among staff and administrators, or successful lobbying for program continuation by interested stakeholders.[5] Whatever the difference, whatever its magnitude, and whomever it affects, this potential for "making a difference" demands defensible evaluation results. For naturalistic evaluation, this defense can be aided, in practically and politically important ways, by the trustworthiness analyses offered by Guba (1981) and Guba and

Lincoln (1981).

Second, the use of such analyses need not necessarily violate the assumptions of naturalistic inquiry. For example, a dependability audit should not attempt to assess whether the methods used were the "best" available, but only that they constitute one set of sound and comprehensive strategies among many such possible sets. A "problem" revealed in a dependability audit is thus more likely to be the omission of data from a relevant program constituency than the choice of a particular interviewing strategy at a particular juncture in the study.

A second example, also cited by Smith (1984), is that of member checks. In his discussion, Smith argues that (a) both the initial inquiry and the member check itself can evoke changes in the subject's perceptions and interpretations of the phenomenon at hand, but there is no definitive way to differentiate among these initial, changed, and re-changed perceptions; and (b) "there can be numerous, yet different, coherent interpretations" (p. 389) of a given phenomenon. From the practical perspective of the practicing evaluator, neither of these concerns is necessarily problematic to his/her use of member checks. As noted above, descriptive member checks focus on the degree to which the evaluator has accurately summarized what was told to, observed by, or reviewed by him/her. And while confirmation of this descriptive account is expected, major discrepancies, if they occur but do not persuade the evaluator to revise his/her account, could be treated like additional data and reported as

such. Also, with the use of interpretive member checks, an evaluator should be fully prepared to hear and make use of a multiplicity of coherent interpretations.

In short, trustworthiness analyses are important instruments in the evaluator's toolkit, designed especially to strengthen the defensibility of results that can make a difference. Further, appropriate use of these instruments need not inevitably undermine the integrity of the naturalistic perspective on meaningfulness in human phenomena.

Qualitative Data Analysis in Practice

Using this framework of general methodological guidance for qualitative data analysis within a naturalistic approach to inquiry, the analysis strategies used in three naturalistic studies are discussed next (and profiled in Table 2).

Insert Table 2 about here

Adult learning study. This exploratory research study was designed, in part, to reveal the nature of participants' learnings in locally initiated adult community groups. Two project team members jointly conducted semi-structured, open-ended interviews with a purposive sample of 10 leaders and members of such groups. The only formal analysis strategy used for this study purpose was an inductive content analysis of the interview transcripts after data collection. This analysis was conducted emergently and iteratively, following the general

naturalistic guidelines described previously. Two features of this analysis, however, warrant special mention (and are the reason why this non-evaluative study is included).

First, the analysis focused on a single, albeit complex, phenomenon, namely, the nature of participant learnings. With this delimited focus, we were able to facilitate significantly our iterative category development with the use of a "proxy list" comprising 114 learning statements (our recording units) initially extracted from the raw data. By using this list as a proxy for the raw data while developing our category system, we were able to work with several pages of text versus 200 single-spaced pages of raw data. Also, by using the raw data for contextual reference during this process and by discarding the proxy list prior to final coding, we attempted to insure that this list did not assume inaccurate meaning or undue importance.

The second noteworthy feature of this analysis was participation by all four project team members. Most importantly, this team approach allowed for numerous analytic discussions during the analysis process. Though unplanned, this team approach also evoked a category development strategy that proved extremely useful to us and thus may be of use in similar contexts. Briefly, this strategy involved the blending or merging of a set of relatively general, abstract categories (e.g., learnings about "self") with a set of more specific, descriptive categories (e.g., "own confidence and perspective strengthened"). The first set was developed collaboratively by three team members who

abstracted "themes" from repeated reviews of the proxy list. The second set was developed independently by the fourth team member, who cut up and sorted all items on the proxy list and then labelled the resulting piles descriptively. A cross-tabulation of the two sets of categories formed the basis for the final category system, which, for illustration purposes, is presented in Table 3. No assessments for trustworthiness were conducted in this study. [More complete discussions of this study's methodology and analysis and its substantive findings can be found in Greene et al. (1984) and Ruiz et al. (1984), respectively.]

Insert Table 3 about here

Program development study. This evaluation study was designed to assess the role of information (its nature and meaning, how it is collected and used, perceived satisfaction with and perceived needs for more) in the program development process of a state level Cooperative Extension (CE) organization. For one component of this study (the other being a statewide mail questionnaire), unstructured, open-ended interviews were conducted with a purposive sample of 27 CE volunteers and staff in two selected counties.

During data collection, interviewers recorded analytic comments on each interview log and, throughout the inquiry, the principal investigator maintained a complete audit trail of all

project activities, decisions, and their rationale. This experience revealed that one somewhat hidden value of maintaining an audit trail is that it encourages consideration of alternatives (for methodological decisions and analytic hypotheses alike) and thereby can strengthen the rationale for the option selected at that time. The mid-stream project team meeting noted in Table 2 was intended as a forum for group analytic induction, specifically for sharing emerging themes and refining interviewing strategies. The meeting, however, was not particularly successful, due largely to time constraints and a lack of structure to the discussion, factors probably also responsible for the overall absence of meaningful analysis during data collection.

Thus, like the adult learning study, the major analysis technique used in this program development study was an inductive content analysis of the interview data after all data were in, again following the general naturalistic guidelines described previously. And in both studies, similar in their short duration, their reliance on interview data, and their absence of meaningful analysis during data collection, inductive content analysis was a useful and appropriate post-data collection strategy. Several differences between the two analyses, however, are also important.

First, in the program development study analysis, we sought a holistic, contextual description and understanding of the nature and role of information, in contrast to the limited focus

on learnings in the adult study. The categories that evolved in the program development study were thus broader and more complex or differentiated (e.g., three "tiers" instead of two -- compare Tables 3 and 4). And the proxy list strategy was not appropriate overall, though it was used and found useful within a couple of the major categories. Second, the program development study analysis was conducted primarily by a single analyst, thus precluding the analytic discussions and the blending of independently-derived category systems that were so beneficial in the adult learning study. Third, however, this lone analyst was able to check her own analytic insights and understandings, as reflected in the category system developed, against the analytic perceptions of others, specifically, the analytic comments interviewers recorded in their logs. That is, though not used effectively during data collection, these comments found an important confirmatory role during post-data collection analysis.

This cross-checking of analytic comments and insights generated independently by different team members is a form of investigator triangulation. More generally, investigator triangulation in qualitative data analysis is a potentially powerful and perhaps underutilized strategy for enhancing the confirmability and credibility of results. Triangulation here means independent, concurrent analyses by two or more investigators, with subsequent assessments of areas of convergence, complementarity, divergence, and difference. While

complete convergence is not expected, given the differing reasoning processes of different analysts, substantial divergence may indicate the presence of strong biases or preconceptions that need to be addressed.

For illustration purposes, the final category system developed in the program development study is shared in Table 4. [See Greene (1984) for a full discussion of the substantive findings.]

Insert Table 4 about here

A draft of the descriptive findings of this study was shared with most interviewees and interviewers by way of member checks. These yielded a number of minor comments, along with strong support for the credibility of the study results. Unfortunately, however, despite the careful maintenance of an audit trail, external audits for confirmability and dependability were not conducted, primarily because funds for such purposes were not available. Internally, the primary analyst did review the "methodological comments" interviewers also recorded in their logs as one check on the quality of the data base. This review indicated, for example, that in all but four interviews, rapport and cooperation were perceived as good and the information exchanged perceived as honest or valid and relevant. However, independent audits of the quality of the data base and the inferences induced from it were again absent.

Stakeholder evaluation study. In this in-progress study, selected assumptions and components of the stakeholder approach to evaluation are being investigated via a case study methodology. Two stakeholder-based evaluations are being conducted in cooperation with two small, local human service agencies. The primary data in this study are thus the qualitative fieldnotes, interviews, meeting minutes, etc. relevant to our understanding of stakeholder evaluation. (Secondary data will come from the evaluations themselves, which include both quantitative and qualitative components and are currently in process.)

In contrast to the short-term, more descriptive nature of the first two studies, this study has a two-and-one-half-year time frame and is intended both to describe and explain. Analysis during the process of data collection is thus more feasible and perhaps more important. As suggested previously, the continual induction and refinement of working hypotheses can contribute significantly to the soundness of final conclusions. Multiple procedural aids to ongoing analytic induction are being employed in this study (see Table 2), all of which are proving useful and important. For example, as part of the audit trail, a methodological log is maintained for each activity conducted with stakeholders. This log includes (a) purpose of the activity; (b) procedures and rationales, including rationales for rejecting alternative procedures; (c) results, including verbatim comments as appropriate; (d) uses of results; and (e) reflections on the

activity, its results, and/or its significance and meaning to the overall study. Beyond these ongoing analysis strategies, this stakeholder evaluation study is valued for its future opportunities to apply some of the lessons learned from previous naturalistic analysis experiences, including the other two shared herein.

Summary

The primary intent of this paper has been to share examples of and reflections on strategies used to analyze qualitative data within the world view represented by the naturalistic approach to inquiry. The spirit of this sharing is to contribute to the "demystification" of such analyses by adding a few more lessons learned to our collective experience, rather than to build toward a codification of qualitative analysis procedures. The stance taken in this paper is that the latter is precluded by the fundamental reliance of qualitative data analysis on the individual analyst's processes of reasoning, understanding, and insight.

Further, this discussion of analysis strategies focused on the particular inquiry contexts presented by program evaluation studies. It was argued that the "real world," potentially impactive nature of evaluation contexts demands broadly defensible results on grounds beyond the individual evaluator's own subjective insights. The trustworthiness analyses offered by Guba (1981) and Guba and Lincoln (1981) were cited as particularly useful for this defense. This argument acknowledged

that such external or independent confirmation of inquiry results is not required by the naturalistic perspective and, as suggested by Smith (1984), is actually epistemologically incompatible with it. Yet, from the pragmatic perspective of the politically attuned evaluator, broadly defensible results, compared to "This is just my view," are more acceptable, and thus more likely to be believed and used. Providing this defense, e.g., through trustworthiness analyses, need not necessarily undermine the integrity of the naturalistic perspective. In fact, this "defensible results strategy" may help enhance the overall acceptability of the naturalistic approach in the public policy and evaluation arena, providing some genuine opportunities for evaluators to demonstrate the value of this perspective on human phenomena.

One final note, again from the view of the practicing evaluator, concerns the resource requirements for naturalistic evaluation. All of the analysis strategies shared herein demand a considerable investment of time by top professional staff, yet resources for evaluation continue to shrink. The response to this squeeze, I believe, is <u>not</u> to compromise on the analytic requirements of valid, defensible inquiry, but rather to conduct smaller studies, with reduced scope and range. This suggestion complements that of others for multiple, even "competing," evaluation studies, smaller in scope than the single "blockbuster" study and differentiated by the designs and methods of separate project teams (Cronbach & Associates, 1980) or by

orientation to a single stakeholder group (Cohen, 1983; Weiss, 1983). This multi-method perspective on program evaluation blends well with the analytic challenges and potential benefits of naturalistic inquiry.

Notes

[1]This linkage, however, refers to practice, not theory. The neutrality of methods, or the absence of an inherent link between methods and philosophical or paradigmatic assumptions, represents one consensus that has emerged from the "quantitative-qualitative" debate (e.g., Bednarz, 1983; Reichardt & Cook, 1979).

[2]Many other labels have been used for this perspective, including interpretive, phenomenological, ethnographic, and qualitative. For convenience and clarity, the "naturalistic" label, as defined and presented by Guba and Lincoln (1981), will be used in this paper.

[3]It should be emphasized that these clusters do not demarcate independent or exclusive turf. Rather, they overlap in that analytic processes described within one cluster or phase of inquiry can and often should be used within another phase.

[4]These are a few examples of a host of naturalistic strategies for insuring that results are confirmable and dependable (see Guba & Lincoln, 1981), not "just the evaluator's own opinion."

[5]These are instances of instrumental, conceptual, and symbolic uses of evaluation results, respectively (Leviton & Hughes, 1982). Naturalistic evaluation results may find their greatest use in the conceptual sphere (McClintock & Greene, 1985).

References

Becker, H.S. (1970). Problems of inference and proof in participant observation. In W.J. Filstead (Ed.). Qualitative methodology. Chicago: Rand McNally.

Bednarz, D. (1983). Quantity and quality in evaluation research: A divergent view. Revised version of paper presented at the Joint Meeting of the Evaluation Research Society and the Evaluation Network, Chicago.

Berelson, B. (1952). Content analysis in communication research. Glencoe, IL: The Free Press.

Bogdan, R.C., & Biklen, S.K. (1982). Qualitative research for education: An introduction to theory and methods. Boston: Allyn and Bacon.

Campbell, D.T. (1979). "Degrees of freedom" and the case study. In T.D. Cook and C.S. Reichardt (Eds.). Qualitative and quantitative methods in evaluation research. Beverly Hills: Sage.

Cohen, D.K. (1983). Evaluation and reform. In A.S. Bryk (Ed.). Stakeholder-based evaluation. Beverly Hills: Sage.

Cronbach, L.J., & Associates (1980). Toward reform of program evaluation. San Francisco: Jossey-Bass.

Daillak, R.H., & Alkin, M.C. (1981). Qualitative studies in context: Reflections on the CSE studies of evaluation use. CSE Report No. 159. Los Angeles: CSE, UCLA.

Denzin, N.K. (1978). The research act: An introduction to sociological methods. New York: McGraw Hill.

Donmoyer, R. (1983). Evaluation as deliberation: Issues, assumptions, and findings from an exploratory study. Final Report (Grant No. NIE-G-81-0083) submitted to the National Institute of Education.

Donmoyer, R. (1984). Choosing from plausible alternatives in interpreting qualitative data. Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans.

Fetterman, D.H. (Ed.). (1984). Ethnography in educational evaluation. Beverly Hills: Sage.

Gillespie, G.W. (1982). Data collection, analysis, and research strategies: A framework for organizing the fieldnote file. Cornell Journal of Social Relations, 16, 98-117.

Glaser, B. (1978). Theoretical sensitivity. Mill Valley, CA: Sociology Press.

Greene, J.C. (1984). Toward enhancing evaluation use: Organizational and methodological perspectives. Paper presented at the Joint Meeting of the Evaluation Research Society and the Evaluation Network, San Francisco.

Greene, J.C., Compton, J.L., Ruiz, E., & Sappington, H. (1984). Some successful strategies for implementing qualitative methods. Paper submitted for publication. Cornell University, Ithaca, NY.

Guba, E.G. (1981). Criteria for assessing the trustworthiness of naturalistic inquiry. Educational Communication and Technology Journal, 29, 75-91.

Guba, E.G., & Lincoln, Y.S. (1981). Effective evaluation. San Francisco: Jossey-Bass.

Holsti, O.R. (1969). Content analysis for the social sciences and humanities. Reading, MA: Addison Wesley.

House, E.R. (1981). Evaluating with validity. Beverly Hills: Sage.

Krippendorf, K. (1980). Content analysis: An introduction to its methodology. Beverly Hills: Sage.

Leviton, C.C., & Hughes, E.F. (1981). Research on the utilization of evaluation: A review and synthesis. Evaluation Review, 5, 525-548.

Lofland, J. (1971). Analyzing social settings: A guide to qualitative observation and analysis. Belmont, CA: Wadsworth.

Marshall, C. (1984). The wrong time for mechanistics in qualitative research. Letter, Educational Researcher, 13, 26-28.

Miles, M.B. (1983). Qualitative data as an attractive nuisance: The problem of analysis. In J. Van Maanen (Ed.). Qualitative methodology. Beverly Hills: Sage.

Miles, M.B., & Huberman, A.M. (1984). Qualitative data analysis: A sourcebook of new methods. Beverly Hills: Sage.

McClintock, C., & Greene, J.C. (1985). Triangulation in practice. Paper submitted for publication. Cornell University, Ithaca, NY.

Patton, M.Q. (1980). Qualitative evaluation methods. Beverly Hills: Sage.

Reichardt, C.S., & Cook, T.D. (1979). Beyond qualitative versus quantitative methods. In T.D. Cook and C.S. Reichardt (Eds.). Qualitative and quantitative methods in evaluation research. Beverly Hills: Sage.

Ruiz, E., Sappington, H., Compton, J.L., & Greene, J.C. (1984). Adult learning through participation in rural community groups. Paper submitted for publication. Cornell University, Ithaca, NY.

Sadler, D.R. (1981). Intuitive data processing as a potential source of bias in naturalistic evaluations. Educational Evaluation and Policy Analysis, 3 (4), 25-31.

Schwandt, T.A. (1984). Fostering disciplined inquiry through an educational audit. Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans.

Smith, J.K. (1984). The problem of criteria for judging interpretive inquiry. Educational Evaluation and Policy Analysis, 6, (4), 379-391.

Smith, J.K. (1983a). Quantitative versus interpretive: The problem of conducting social inquiry. In E.R. House (Ed.). Philosophy of evaluation. San Francisco: Jossey-Bass.

Smith, J.K. (1983b). Quantitative versus qualitative research: An attempt to clarify the issue. Educational Researcher, 12, 6-13.

Smith, J.K., & Heshusius, L. (1985). Closing down the conversation: The end of the quantitative-qualitative debate among educational inquirers. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago.

Weiss, C.H. (1983). Toward the future of stakeholder approaches in evaluation. In A.S. Bryk (Ed.). Stakeholder-based evaluation. Beverly Hills: Sage.

Table 1

Content Analysis and Inductive Content Analysis: Similarities and Differences

Similarities	Differences	
	Content analysis	Inductive content analysis
Designed for use with large volumes of qualitative information	A research strategy	A data analysis technique
Intended to describe and/or explain phenomena with contextual sensitivity	Designed for use with existing information, collected or generated for reasons totally unrelated to the inquiry	Designed for use with both existing information and information collected as part of the inquiry
Is a systematic, rule-guided process	Usually intended to describe/explain symbolic content	Usually intended to describe/explain manifest content
Requires identification of units of analysis, both recording and context units	Categories are established deductively, prior to their application in data coding	Categories are established inductively through repeated iterations of category development and data coding
Uses sets of categories that: a. reflect the purpose of the study b. are exhaustive c. are mutually exclusive d. are independent e. reflect a single classification principle	Categories thus represent a priori concepts and meanings assigned by the inquirer	Categories thus represent emergent concepts and meanings of the phenomena being studied
	So categories serve as variables	So categories serve to provide an organizational framework for summarizing the raw data
	And data of major interest are thus the codes	And data of major interest are thus the content of the categories
	Historically, a quantitative strategy that includes inferential statistical analysis of data codes	Largely, a nonquantitative technique (though frequencies or even crosstabs of data codes are sometimes included)

Table 2

Profiles of the Analysis Components of the Three Studies

Analysis component	Adult learning study	Program development study	Stakeholder evaluation study
During data collection	Informal, unrecorded analytic conversations between the two interviewers	Analytic comments included on each interview log Maintenance of audit trail One project team analytic meeting to discuss emerging themes	Analytic comments included on each data log Periodic and inspirational analytic memos Occasional project team analytic discussions of emerging themes and hypotheses Frequent member (stakeholder) checks Maintenance of an audit trail
After data collection	Inductive content analysis, including: *use of a proxy list *blending of two category systems *participation by four project team members	Inductive content analysis conducted primarily by single analyst Check of resulting categories and findings against interviewers' analytic comments	(N/A at this time)
For assessing trustworthiness of results		Internal review of part of the audit trail Member checks	(N/A at this time)

Table 3

Final Category System in the Adult Learning Study

Learnings linked to group participation
Self-related
Own confidence and perspective strengthened
Own capacities and limitations clarified
Personal satisfaction from participation important
Internal knowledge about group processes or lessons learned about...
Patience and tolerance/realistic expectations in group work
Providing leadership
Decision making in groups
Doing successful organizational work
Internal skills regarding groups or how to ...
Get and keep people involved
Lead groups
Set up and run an organization
External knowledge related to group work
Knowing the community, relying on local resources/initiatives
Gathering information
Being organized
Understanding the dynamics of power and the process of change
Recognizing apathy and nonparticipation
Knowing about government and the media
Networking with other individuals/groups
External skills regarding groups or how to...
Carry out organizational tasks
Deal with government
Broader skills extending beyond groups or how to...
Get along with different kinds of people
Be assertive/take risks
Be a leader
Evaluate and be objective
Knowledge about special groups and issues
Special age groups (the elderly, children)
Minorities
Environment and energy
Learning about learning

Table 4

Final Category System in the Program Development Study

Components of the program development process related to the role of information
People factors
Personal history of CE staff and volunteers
Length of time and roles in CE
Committment to and benefits from CE
Personal "connections" and their benefits and limitations
Social connections
Political connections
Skills, knowledge, expertise of CE people
Intraorganizational perceptions of CE people, including trust factor
Links between staff expertise and program development
Organizational factors
Roles, responsibilities, and interrelationships of CE people as related to program development
Roles, responsibilities of program committees
Roles, responsibilities of Board
Roles, responsibilities of county staff
Roles, responsibilities of state staff and faculty
Interrelationships of Board with other groups
Interrelationships of program committees with county staff
Interrelationships of county staff with state staff and faculty
Networking with other CE groups and organizations outside CE
Contextual climate for CE
Overall perceptions of CE and CE programs, including reputation, visibility, trust factor
Perceptions of purpose, mission of CE
Perceptions of general nature of CE planning (bottom up)
Organizational needs
Funding priorities
In-service needs
Needs for new or expanded programs and staff
Changes needed in general planning

Table 4 (continued)

Components of the program development process related to the role of information
Information
Existing, formal information sources
Records
Program committee minutes
Plans
Reports
Information used in the program development process
Census, demographic information
Local needs information
Information from the university and faculty research
Information about staff (distribution of time, expertise)
Information from staff
Information re previous activities (attendance, phone calls, waiting lists)
More information re previous activities (informal perceptions or feedback from participants)
Information from evaluations
Specific instances of information and information sources
Information needed in the program development process
No, none
More input
More feedback
Program development processes
Long-range planning and review processes
Developing/generating program ideas
Determining/assessing program needs, priorities, target audiences
Developing and implementing programs
Evaluating and reporting on programs
Information flow and exchange
Improvements needed